SEQUENCE LISTING

<110> Ingram, L et al.

<120> METHODS AND COMPOSITIONS FOR SIMULTANEOUS
      SACCHARIFICATION AND FERMENTATION

<130> BCI-024CP

<140>
<141>

<150> 60/214,137
<151> 2000-06-26

<150> 60/219,913
<151> 2000-07-21

<160> 17

<170> PatentIn Ver. 2.0

<210> 1
<211> 450
<212> DNA
<213> Zymomonas mobilis

<220>
<223> promoter

<400> 1
cttttttcggc atgagcaacc aacatttttca aggtatcatc ctgatgcgca atatcggcat 60

cggttagcca taaccattt acctgtccgg cggccttaat accttgatca gatggttcgt 120

ggtgttgtta ccttgccgaa gggcaccggt aaaaatgttc gcgtcggtgt tttcgcccgt 180

ggcccgaaag ctgaagaagc taaagctgct ggtgcagaag ttgtcggcgc agaagacctg 240

atggaagcca ttcagggcgg cagcattgat ttcgatcgtg atgccctta tactgaaatt 300

gccttgcgct gccataatga agcagcctcc ggtgttttgg cagatttaag cgctgcctga 360

ttttcgtgat cctctagagt ctatgaaatg gagattcatt tatgcctctc tcttattcgg 420

ataaccatcc agtcatccgc aagcttggcc   450

<210> 2
<211> 1509
<212> DNA
<213> Zymomonas mobilis

<220>
<223> expression
      vector

<400> 2
gatcaaccgg caatttatcc acggcatcaa attcgatctg tcttttcccg tatcattggc 60

aataccggca ttctgattac aggccgtgtt ttgaatgcgg tatgcagttt tgtctatgtc 120

```
gcatggacat cccagacatt gggattgaac ctgtttggtg tcatgctttt gattacgact  180

tttgctaccc tgatttcgga tattacccgt tttcagtcat ggcaaacctt gctgcattac  240

ggttcaaaag cttttcagga aaaagatttt aaccaatttg atgatgtcct tgccttttgc  300

atcagagccg attttttttag tgcggcgata ggtatgttgg tagggttagg cggtatcttg  360

attttaggca cttcaagatt gggatggcct gccgaggtca agccagatgc cttgctttgt  420

atgctgatta tactttttat gaatatcggc tggtccaacc gggatgttgc ggctgtgtaa  480

ccgctttaaa ctggtcacta tttatgagtt tattacgacc tgcgtcagaa ccggaggttg  540

tggcattggt tattggcttc atatgccttt ggggtatttt ttgtttatat ggtgcctgac  600

gcaattcacg ctttttgtca cctgtagtta cgctggcatt tatctctttc accaatatac  660

ggagcgagca tttccgataa gaaaaatatt tcagagaaaa acgcccgttg aagggatgtg  720

gaaattcact ttaagcgtca gtttttaatga aatcctagac tccattttcc agcagggtgg  780

caccccttgct attggtagct cactgggggc tggggaagcc gctgtctatc gggtcgcgcg  840

ccagattagt aacggtttat ccaaaccagc acagatgatg atcggctaac atgcatccac  900

cggcagcacc ggccgtttta tgcttgggat tattgatatg ccgaaaagga tacaacatct  960

ggaagaaaaa gacgaaggcc ggaataagcg cccattctgc aaaattgtta caacttagtc 1020

gcgccatcag ggaatgaaaa atcaatccgt cttttttcggc atgagcaacc aacattttca 1080

aggtatcatc ctgatgcgca atatcggcat cggttagcca taaccatttt acctgtccgg 1140

cggccttaat accttgatca gatggttcgt ggtgttgtta ccttgccgaa gggcaccggt 1200

aaaaatgttc gcgtcggtgt tttcgcccgt ggcccgaaag ctgaagaagc taaagctgct 1260

ggtgcagaag ttgtcggcgc agaagacctg atggaagcca ttcagggcgg cagcattgat 1320

ttcgatcgtg atgcccttta tactgaaatt gccttgcgct gccataatga agcagcctcc 1380

ggtgttttgg cagatttaag cgctgcctga ttttcgtgat cctctagagt ctatgaaatg 1440

gagattcatt tatgcctctc tcttattcgg ataaccatcc agtcatccgc aagcttggcc 1500

gtaatccat                                                         1509
```

<210> 3
<211> 27
<212> DNA
<213> Escherichia coli

<220>
<223> Description of Artificial Sequence:primer

<400> 3
cgaattcctg ccgaagttta ttagcca   27

```
<210> 4
<211> 31
<212> DNA
<213> Escherichia coli

<220>
<223> Description of Artificial Sequence:primer

<400> 4
aaggatcctt ccaccagcta tttgttagtg a                                31


<210> 5
<211> 27
<212> DNA
<213> Escherichia coli

<220>
<223> Description of Artificial Sequence:primer

<400> 5
agaattctgc cagttggttg acgatag                                    27


<210> 6
<211> 30
<212> DNA
<213> Escherichia coli

<220>
<223> Description of Artificial Sequence:primer

<400> 6
caggatcccc tcaagtcact agttaaactg                                 30


<210> 7
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 7
taatacgact cactataggg                                            20


<210> 8
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 8
taacaatttc acacagga                                              18
```

```
<210> 9
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 9
cacgacgttg taaaacgac                                                    19


<210> 10
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 10
gactggatgg ttatccgaat aagagagagg                                       30


<210> 11
<211> 8
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 11
ggcgcgcc                                                                8


<210> 12
<211> 11544
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:vector

<220>
<223> all occurrences of n to be sequenced

<220>
<223> nucleotide positions 1-1451 ecodes promoter

<220>
<223> nucleotide positions 1452-2735 encodes celZ gene

<220>
<223> nucleotide positions 4916-5776 encodes bla gene

<220>
<223> nucleotide positions 7061-8251 encodes tet gene

<220>
<223> nucleotide positions 9476-11544 encodes target sequence from K. oxytoca
```

<220>
<221> CDS
<222> (1452)..(2735)

<220>
<221> CDS
<222> (4916)..(5776)

<220>
<221> CDS
<222> (7061)..(8251)

<400> 12
gatcaaccgg caatttatcc acggcatcaa attcgatctg tcttttcccg tatcattggc 60

aataccggca ttctgattac aggccgtgtt ttgaatgcgg tatgcagttt tgtctatgtc 120

gcatggacat cccagacatt gggattgaac ctgtttggtg tcatgctttt gattacgact 180

tttgctaccc tgatttcgga tattacccgt tttcagtcat ggcaaacctt gctgcattac 240

ggttcaaaag cttttcagga aaaagatttt aaccaatttg atgatgtcct tgccttttgc 300

atcagagccg attttttttag tgcggcgata ggtatgttgg tagggttagg cggtatcttg 360

attttaggca cttcaagatt gggatggcct gccgaggtca agccagatgc cttgctttgt 420

atgctgatta tactttttat gaatatcggc tggtccaacc gggatgttgc ggctgtgtaa 480

ccgctttaaa ctggtcacta tttatgagtt tattacgacc tgcgtcagaa ccggaggttg 540

tggcattggt tattggcttc atatgccttt ggggtatttt ttgtttatat ggtgcctgac 600

gcaattcacg ctttttgtca cctgtagtta cgctggcatt tatctctttc accaatatac 660

ggagcgagca tttccgataa gaaaaatatt tcagagaaaa cgcccgttg aagggatgtg 720

gaaattcact ttaagcgtca gttttaatga aatcctagac tccattttcc agcagggtgg 780

caccccttgct attggtagct cactgggggc tggggaagcc gctgtctatc gggtcgcgcg 840

ccagattagt aacggtttat ccaaaccagc acagatgatg atcggctaac atgcatccac 900

cggcagcacc ggccgtttta tgcttgggat tattgatatg ccgaaaagga tacaacatct 960

ggaagaaaaa gacgaaggcc ggaataagcg cccattctgc aaaattgtta caacttagtc 1020

gcgccatcag ggaatgaaaa atcaatccgt cttttcggc atgagcaacc aacattttca 1080

aggtatcatc ctgatgcgca atatcggcat cggttagcca taaccatttt acctgtccgg 1140

cggccttaat accttgatca gatggttcgt ggtgttgtta ccttgccgaa gggcaccggt 1200

aaaaatgttc gcgtcggtgt tttcgcccgt ggcccgaaag ctgaagaagc taaagctgct 1260

ggtgcagaag ttgtcggcgc agaagacctg atggaagcca ttcagggcgg cagcattgat 1320

ttcgatcgtg atgcccttta tactgaaatt gccttgcgct gccataatga agcagcctcc 1380

ggtgtttttgg cagatttaag cgctgcctga ttttcgtgat cctctagagt ctatgaaatg 1440

```
gagattcatt t atg cct ctc tct tat tcg gat aac cat cca gtc atc gat    1490
           Met Pro Leu Ser Tyr Ser Asp Asn His Pro Val Ile Asp
            1               5                   10

agc caa aaa cac gcc cca cgt aaa aaa ctg ttt cta tct tgt gcc tgt    1538
Ser Gln Lys His Ala Pro Arg Lys Lys Leu Phe Leu Ser Cys Ala Cys
     15                  20                  25

tta gga tta agc ctt gcc tgc ctt tcc agt aat gcc tgg gcg agt gtt    1586
Leu Gly Leu Ser Leu Ala Cys Leu Ser Ser Asn Ala Trp Ala Ser Val
 30                  35                  40                  45

gag ccg tta tcc gtt agc ggc aat aaa atc tac gca ggt gaa aaa gcc    1634
Glu Pro Leu Ser Val Ser Gly Asn Lys Ile Tyr Ala Gly Glu Lys Ala
                 50                  55                  60

aaa agt ttt gcc ggc aac agc tta ttc tgg agt aat aat ggt tgg ggt    1682
Lys Ser Phe Ala Gly Asn Ser Leu Phe Trp Ser Asn Asn Gly Trp Gly
             65                  70                  75

ggg gaa aaa ttc tac aca gcc gat acc gtt gcg tcg ctg aaa aaa gac    1730
Gly Glu Lys Phe Tyr Thr Ala Asp Thr Val Ala Ser Leu Lys Lys Asp
             80                  85                  90

tgg aaa tcc agc att gtt cgc gcc gct atg ggc gtt cag gaa agc ggt    1778
Trp Lys Ser Ser Ile Val Arg Ala Ala Met Gly Val Gln Glu Ser Gly
     95                  100                 105

ggt tat ctg cag gac ccg gct ggc aac aag gcc aaa gtt gaa aga gtg    1826
Gly Tyr Leu Gln Asp Pro Ala Gly Asn Lys Ala Lys Val Glu Arg Val
110                 115                 120                 125

gtg gat gcc gca atc gcc aac gat atg tat gtg att att gac tgg cac    1874
Val Asp Ala Ala Ile Ala Asn Asp Met Tyr Val Ile Ile Asp Trp His
                 130                 135                 140

tca cat tct gca gaa aac aat cgc agt gaa gcc att cgc ttc ttc cag    1922
Ser His Ser Ala Glu Asn Asn Arg Ser Glu Ala Ile Arg Phe Phe Gln
             145                 150                 155

gaa atg gcg cgc aaa tat ggc aac aag ccg aat gtc att tat gaa atc    1970
Glu Met Ala Arg Lys Tyr Gly Asn Lys Pro Asn Val Ile Tyr Glu Ile
         160                 165                 170

tac aac gag ccg ctt cag gtt tca tgg agc aat acc att aaa cct tat    2018
Tyr Asn Glu Pro Leu Gln Val Ser Trp Ser Asn Thr Ile Lys Pro Tyr
     175                 180                 185

gcc gaa gcc gtg att tcc gcc att cgc gcc att gac ccg gat aac ctg    2066
Ala Glu Ala Val Ile Ser Ala Ile Arg Ala Ile Asp Pro Asp Asn Leu
190                 195                 200                 205

att att gtc ggt acg ccc agt tgg tcg caa aac gtt gat gaa gcg tcg    2114
Ile Ile Val Gly Thr Pro Ser Trp Ser Gln Asn Val Asp Glu Ala Ser
                 210                 215                 220

cgc gat cca atc aac gcc aag aat atc gcc tat acg ctg cat ttc tac    2162
Arg Asp Pro Ile Asn Ala Lys Asn Ile Ala Tyr Thr Leu His Phe Tyr
             225                 230                 235
```

```
gcg gga acc cat ggt gag tca tta cgc act aaa gcc cgc cag gcg tta    2210
Ala Gly Thr His Gly Glu Ser Leu Arg Thr Lys Ala Arg Gln Ala Leu
        240                 245                 250

aat aac ggt att gcg ctt ttc gtc acc gag tgg ggc gcc gtt aac gcg    2258
Asn Asn Gly Ile Ala Leu Phe Val Thr Glu Trp Gly Ala Val Asn Ala
    255                 260                 265

gac ggc aat ggc gga gtg aac cag aca gat acc gac gcc tgg gta acg    2306
Asp Gly Asn Gly Gly Val Asn Gln Thr Asp Thr Asp Ala Trp Val Thr
270                 275                 280                 285

ttc atg cgt gac aac aac atc agc aac gca aac tgg gcg tta aat gat    2354
Phe Met Arg Asp Asn Asn Ile Ser Asn Ala Asn Trp Ala Leu Asn Asp
                290                 295                 300

aaa agc gaa ggg gca tca acc tat tat ccg gac tct aaa aac ctg acc    2402
Lys Ser Glu Gly Ala Ser Thr Tyr Tyr Pro Asp Ser Lys Asn Leu Thr
                305                 310                 315

gag tcg ggt aaa ata gta aaa tcg atc att caa agc tgg cca tat aaa    2450
Glu Ser Gly Lys Ile Val Lys Ser Ile Ile Gln Ser Trp Pro Tyr Lys
        320                 325                 330

gcg ggc agc gcc gcc agt aca aca acc gat cag tca acc gat acc acc    2498
Ala Gly Ser Ala Ala Ser Thr Thr Thr Asp Gln Ser Thr Asp Thr Thr
        335                 340                 345

atg gca cca ccg ttg acg aac cga cca caa ccg aca cac cgg caa acc    2546
Met Ala Pro Pro Leu Thr Asn Arg Pro Gln Pro Thr His Arg Gln Thr
350                 355                 360                 365

gct gat tgc tgc aat gcc aac gtt tac ccc aac tgg gtt agc aaa gac    2594
Ala Asp Cys Cys Asn Ala Asn Val Tyr Pro Asn Trp Val Ser Lys Asp
                370                 375                 380

tgg gcg ggc cgg cag cga ctc ata acg aag cag gcc aat cga tcg tct    2642
Trp Ala Gly Arg Gln Arg Leu Ile Thr Lys Gln Ala Asn Arg Ser Ser
        385                 390                 395

aca aag gga acc tgt ata ccg caa act ggt aca ctt cat ccg ttc cgg    2690
Thr Lys Gly Thr Cys Ile Pro Gln Thr Gly Thr Leu His Pro Phe Arg
        400                 405                 410

gca gcg att cct cct ggg cac agg ttg gta gct gta act aat tga        2735
Ala Ala Ile Pro Pro Gly His Arg Leu Val Ala Val Thr Asn
        415                 420                 425

ttaatctttt cacccccaaa ataacagggc tgcgattgca gcctgatacg caacattcca  2795

ttacttaatt gcgttcaaaa gcgcccaaat ccggtgcgct gccttgtaac taatatgatt  2855

tctctttcgt acccgcgtta atcagctttg agttagccga cagacggaac agcgaggttg  2915

ccggcaacgt gccgtcatta tcacgagata cggtagccag cgaggtgtcc aggctgacga  2975

atcggacgcg gaagccgctg tccgtatcca tgagttgact cgcatccgca ttactgaccg  3035

ttgcagaagc agacagagac acgttgttgc ggaagtaatg tttctgtcct gactggacgt  3095

tgctcccgaa agcataatta atgccgtttt tatatgacgt gttatttatt accgtacgcc  3155
```

```
gccgcgttat tgttctggtc aaaacctttg ctcacgttgc caaacgcgac gcaacgggta 3215

atgcgatgat tgccgaccgc tggttcctcc cagtttgaac ccgttggcat tgccggcgaa 3275

cgcgctnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3335

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3395

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3455

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3515

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3575

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3635

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3695

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3755

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3815

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3875

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3935

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3995

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4055

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4115

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4175

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4235

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnngatc ctctagagtc 4295

gacctgcagg aattcactgg ccgtcgtttt acaacgtcgt gactgggaaa accctggcgt 4355

tacccaactt aatcgccttg cagcacatcc ccctttcgcc agctggcgta atagcgaaga 4415

ggcccgcacc gatcgccctt cccaacagtt gcgcagcctg aatggcgaat ggcgcctgat 4475

gcggtatttt ctccttacgc atctgtgcgg tatttcacac cgcataggcg cgcctatggt 4535

gcactctcag tacaatctgc tctgatgccg catagttaag ccagccccga cacccgccaa 4595

cacccgctga cgcgccctga cgggcttgtc tgctcccggc atccgcttac agacaagctg 4655

tgaccgtctc cgggagctgc atgtgtcaga ggttttcacc gtcatcaccg aaacgcgcga 4715

gacgaaaggg cctcgtgata cgcctatttt tataggttaa tgtcatgata ataatggttt 4775

cttagacgtc aggtggcact tttcggggaa atgtgcgcgg aacccctatt tgtttatttt 4835

tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat 4895

aatattgaaa aaggaagagt atg agt att caa cat ttc cgt gtc gcc ctt att 4948
                       Met Ser Ile Gln His Phe Arg Val Ala Leu Ile
```

```
ccc ttt ttt gcg gca ttt tgc ctt cct gtt ttt gct cac cca gaa acg    4996
Pro Phe Phe Ala Ala Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr
                             20

ctg gtg aaa gta aaa gat gct gaa gat cag ttg ggt gca cga gtg ggt    5044
Leu Val Lys Val Lys Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly
         30                                      40

tac atc gaa ctg gat ctc aac agc ggt aag atc ctt gag agt ttt cgc    5092
Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg
                         50

ccc gaa gaa cgt ttt cca atg atg agc act ttt aaa gtt ctg cta tgt    5140
Pro Glu Glu Arg Phe Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys
 60                                      70

ggc gcg gta tta tcc cgt att gac gcc ggg caa gag caa ctc ggt cgc    5188
Gly Ala Val Leu Ser Arg Ile Asp Ala Gly Gln Glu Gln Leu Gly Arg
                 80                                      90

cgc ata cac tat tct cag aat gac ttg gtt gag tac tca cca gtc aca    5236
Arg Ile His Tyr Ser Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr
                             100

gaa aag cat ctt acg gat ggc atg aca gta aga gaa tta tgc agt gct    5284
Glu Lys His Leu Thr Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala
         110                                     120

gcc ata acc atg agt gat aac act gcg gcc aac tta ctt ctg aca acg    5332
Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr
                         130

atc gga gga ccg aag gag cta acc gct ttt ttg cac aac atg ggg gat    5380
Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp
140                                      150

cat gta act cgc ctt gat cgt tgg gaa ccg gag ctg aat gaa gcc ata    5428
His Val Thr Arg Leu Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile
                 160                                     170

cca aac gac gag cgt gac acc acg atg cct gta gca atg gca aca acg    5476
Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Val Ala Met Ala Thr Thr
                             180

ttg cgc aaa cta tta act ggc gaa cta ctt act cta gct tcc cgg caa    5524
Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln
         190                                     100

caa tta ata gac tgg atg gag gcg gat aaa gtt gca gga cca ctt ctg    5572
Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu
                     110

cgc tcg gcc ctt ccg gct ggc tgg ttt att gct gat aaa tct gga gcc    5620
Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala
120                                      130

ggt gag cgt ggg tct cgc ggt atc att gca gca ctg ggg cca gat ggt    5668
Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly
                 140                                     150
```

```
aag ccc tcc cgt atc gta gtt atc tac acg acg ggg agt cag gca act   5716
Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr
                            160

atg gat gaa cga aat aga cag atc gct gag ata ggt gcc tca ctg att   5764
Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile
        170                             180

aag cat tgg taa ctgtcagacc aagtttactc atatatactt tagattgatt      5816
Lys His Trp
    185

taaaacttca tttttaattt aaaaggatct aggtgaagat ccttttgat aatctcatga  5876

ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta gaaaagatca  5936

aaggatcttc ttgagatcct tttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac   5996

caccgctacc agcggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg   6056

taactggctt cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag   6116

gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac   6176

cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt   6236

taccggataa ggcgcagcgg tcgggctgaa cggggggttc gtgcacacag cccagcttgg   6296

agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcgccacgc   6356

ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga acaggagagc   6416

gcacgaggga gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc   6476

acctctgact tgagcgtcga ttttgtgat gctcgtcagg ggggcggagc ctatggaaaa   6536

acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt gctcacatgt   6596

tctttcctgc gttatcccct gattctgtgg ataaccgtat taccgccttt gagtgagctg   6656

ataccgctcg ccgcagccga cgaccgagc gcagcgagtc agtgagcgag gaagcggcgc    6716

gccagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta   6776

atgcagctgg cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa   6836

tgtgagttag ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat   6896

gttgtgtgga attgtgagcg ataacaatt tcacacagga aacagctatg accatgatta   6956

cgccaagctt gcatgccaat tctcatgttt gacagcttat catcgataag ctttaatgcg   7016

gtagtttatc acagttaaat tgctaacgca gtcaggcacc gtgt atg aaa tct aac   7072
                                               Met Lys Ser Asn

aat gcg ctc atc gtc atc ctc ggc acc gtc acc ctg gat gct gta ggc   7120
Asn Ala Leu Ile Val Ile Leu Gly Thr Val Thr Leu Asp Ala Val Gly
                        10                              20

ata ggc ttg gtt atg ccg gta ctg ccg ggc ctc ttg cgg gat atc gtc   7168
```

Ile Gly Leu Val Met Pro Val Leu Pro Gly Leu Leu Arg Asp Ile Val
                                        30

cat tcc gac agc atc gcc agt cac tat ggc gtg ctg cta gcg cta tat   7216
His Ser Asp Ser Ile Ala Ser His Tyr Gly Val Leu Leu Ala Leu Tyr
         40   50

gcg ttg atg caa ttt cta tgc gca ccc gtt ctc gga gca ctg tcc gac   7264
Ala Leu Met Gln Phe Leu Cys Ala Pro Val Leu Gly Ala Leu Ser Asp
                        60

cgc ttt ggc cgc cgc cca gtc ctg ctc gct tcg cta ctt gga gcc act   7312
Arg Phe Gly Arg Arg Pro Val Leu Leu Ala Ser Leu Leu Gly Ala Thr
    70   80

atc gac tac gcg atc atg gcg acc aca ccc gtc ctg tgg atc ctc tac   7360
Ile Asp Tyr Ala Ile Met Ala Thr Thr Pro Val Leu Trp Ile Leu Tyr
                        90   100

gcc gga cgc atc gtg gcc ggc atc acc ggc gcc aca ggt gcg gtt gct   7408
Ala Gly Arg Ile Val Ala Gly Ile Thr Gly Ala Thr Gly Ala Val Ala
                                        110

ggc gcc tat atc gcc gac atc acc gat ggg gaa gat cgg gct cgc cac   7456
Gly Ala Tyr Ile Ala Asp Ile Thr Asp Gly Glu Asp Arg Ala Arg His
            120   130

ttc ggg ctc atg agc gct tgt ttc ggc gtg ggt atg gtg gca ggc ccc   7504
Phe Gly Leu Met Ser Ala Cys Phe Gly Val Gly Met Val Ala Gly Pro
                        140

gtg gcc ggg gga ctg ttg ggc gcc atc tcc ttg cat gca cca ttc ctt   7552
Val Ala Gly Gly Leu Leu Gly Ala Ile Ser Leu His Ala Pro Phe Leu
    150   160

gcg gcg gcg gtg ctc aac ggc ctc aac cta cta ctg ggc tgc ttc cta   7600
Ala Ala Ala Val Leu Asn Gly Leu Asn Leu Leu Leu Gly Cys Phe Leu
                        170   180

atg cag gag tcg cat aag gga gag cgt cga ccg atg ccc ttg aga gcc   7648
Met Gln Glu Ser His Lys Gly Glu Arg Arg Pro Met Pro Leu Arg Ala
                                        190

ttc aac cca gtc agc tcc ttc cgg tgg gcg cgg ggc atg act atc gtc   7696
Phe Asn Pro Val Ser Ser Phe Arg Trp Ala Arg Gly Met Thr Ile Val
            200   210

gcc gca ctt atg act gtc ttc ttt atc atg caa ctc gta gga cag gtg   7744
Ala Ala Leu Met Thr Val Phe Phe Ile Met Gln Leu Val Gly Gln Val
                        220

ccg gca gcg ctc tgg gtc att ttc ggc gag gac cgc ttt cgc tgg agc   7792
Pro Ala Ala Leu Trp Val Ile Phe Gly Glu Asp Arg Phe Arg Trp Ser
    230   240

gcg acg atg atc ggc ctg tcg ctt gcg gta ttc gga atc ttg cac gcc   7840
Ala Thr Met Ile Gly Leu Ser Leu Ala Val Phe Gly Ile Leu His Ala
                        250   260

ctc gct caa gcc ttc gtc act ggt ccc gcc acc aaa cgt ttc ggc gag   7888
Leu Ala Gln Ala Phe Val Thr Gly Pro Ala Thr Lys Arg Phe Gly Glu

270

```
aag cag gcc att atc gcc ggc atg gcg gcc gac gcg ctg ggc tac gtc   7936
Lys Gln Ala Ile Ile Ala Gly Met Ala Ala Asp Ala Leu Gly Tyr Val
            180                                     190

ttg ctg gcg ttc gcg acg cga ggc tgg atg gcc ttc ccc att atg att   7984
Leu Leu Ala Phe Ala Thr Arg Gly Trp Met Ala Phe Pro Ile Met Ile
                        200

ctt ctc gct tcc ggc ggc atc ggg atg ccc gcg ttg cag gcc atg ctg   8032
Leu Leu Ala Ser Gly Gly Ile Gly Met Pro Ala Leu Gln Ala Met Leu
        210                                 220

tcc agg cag gta gat gac gac cat cag gga cag ctt caa gga tcg ctc   8080
Ser Arg Gln Val Asp Asp Asp His Gln Gly Gln Leu Gln Gly Ser Leu
                    230                                 240

gcg gct ctt acc agc cta act tcg atc act gga ccg ctg atc gtc acg   8128
Ala Ala Leu Thr Ser Leu Thr Ser Ile Thr Gly Pro Leu Ile Val Thr
                                    250

gcg att tat gcc gcc tcg gcg agc aca tgg aac ggg ttg gca tgg att   8176
Ala Ile Tyr Ala Ala Ser Ala Ser Thr Trp Asn Gly Leu Ala Trp Ile
            260                                     270

gta ggc gcc gcc cta tac ctt gtc tgc ctc ccc gcg ttg cgt cgc ggt   8224
Val Gly Ala Ala Leu Tyr Leu Val Cys Leu Pro Ala Leu Arg Arg Gly
                        280

gca tgg agc cgg gcc acc tcg acc tga atggaagccg gcggcacctc        8271
Ala Trp Ser Arg Ala Thr Ser Thr
        290
```

gctaacggat tcaccactcc aagaattgga gccaatcaat tcttgcggag aactgtgaat 8331

gcgcaaacca acccttggca gaacatatcc atcgcgtccg ccatctccag cagccgcacg 8391

cggcgcatct cggggtcgac tctagaggat ccccgcaacg ctgtcagcgc tttccagtta 8451

aacggctcca acgtcgccat aggtaattcc tcgcccggcc atacgatcgg gcaggtgccg 8511

ttggctatcg ccgtcgcctg actcatcaca ctatcttccg ctgcatcgcg aagggttttg 8571

accacttctt ccatctctcc gtgcgccgga tgccatgctc acgtacgcgg cttatcagat 8631

agtcgggcag gccgtcgttc cagcccaatg aggggaagct ggcgtggagc gatgccagca 8691

cctgctcctc aacaccgtaa tggccggcgg cgaacaggca ttcggcggta agcgcttcca 8751

gccctttaat catcacgctg cggcacatct tgatagccga cacgctgcca acgtggttac 8811

caccatagcg ggcgttacat ccaagcgtgg tgagtaattc agcaattgcc tctgcctgtg 8871

gtccccccgt caacagcggc gttcggagtg cccctggggg gaccggcgcc atcaccgcta 8931

catcgacata agcgccgggc ttaaagcatt tggcagcctg acgcttggtc tgcggggcga 8991

cggagttaag gtcaagaaaa tactgcgtgt cggtcatcag cggtgcagct tgtgaggcga 9051

catccagggc ggatcccgcg gtgacggtgg aaaatatgag ttcggcacct gtcaacgcgt 9111

```
cagccaggga gattgccgcc cgcacgcctc cccgatgcgc cttcgttatc atcgcatcgc 9171

gctcaggacc ttgcagcttg caatcccaga cggtgactgg gttcactttt gccagtgcat 9231

ccgcaagaat gccacctgct tcaccataac ctataaacgt tattgtcgtc ataacagctc 9291

cttacgcggc cacacgtcgg ccggaatgca aacgtcgccc gcgaacagaa gtcgcgccgt 9351

acgcagcaga ccgcagcctg ccaactgccc attatcatca gccggagcg ccacgctgaa 9411

ttgggtaccg agctccgaat tgggtaccga gctcgaatta attcgagctc ggtacccggg 9471

gatcnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9531

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9591

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9651

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9711

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9771

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9831

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9891

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9951

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10011

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10071

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10131

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10191

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10251

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10311

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10371

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10431

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10491

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10551

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10611

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10671

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10731

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10791

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10851

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10911
```

```
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10971

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11031

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11091

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11151

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11211

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11271

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11331

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11391

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11451

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 11511

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnn                              11544
```

<170> PatentIn Ver. 2.0

<210> 13
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 13
ctgttccgtt accaacac   18

<210> 14
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 14
gtgaatggga tcacgagt   18

<210> 15
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 15
accatcagca tcaacgccca acaacg   26

<210> 16
<211> 30
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 16
gactggatgg ttatccgaat aagagagagg   30

<210> 17
<211> 11772
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
       construct

<220>
<223> unsequenced Erwinia DNA fragment from nucleotide
       position 1143 to 144

<220>
<223> P1 promoter region for celZ from nucleotide
       position 4424 to 2974

<220>
<223> guide fragment for integration from nucleotide
       position 4677 to 7573

<220>
<223> sequenced partial guide fragment from nucleotide
       position 4677 to 5752

<220>
<223> unsequenced partial guide fragment from nucleotide
       position 5753 to 7573

<220>
<223> P2 promoter region for celY from nucleotide
       position 7585 to 8576

<220>
<223> R6K-Y ori from nucleotide position 10388 to 10763

<220>
<223> FRTF lipase-binding sequence from nucleotide
       position 16 to 50

<220>
<223> FRTFlipase-binding sequence from nucleotide
       position 10058 to 10092

<220>
<223> celZ gene product is encoded by the complement of
       nucleotides 2973 to 1690

<220>
<223> celY gene product is encoded by the nucleotides
       8576 to 9574

<220>

<223> kanamicin-resistance gene product is encoded by
      the complement of nucleotides 11621 to 10827

<400> 17
```
atcgatgaat tgatccgaag ttcctattct ctagaaagta taggaacttc gaattgtcga 60
caagcttgat ctggcttatc gaaattaata cgactcacta tagggagacc ggaattcccc 120
tgcaggtcga ctctagagga tcannnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 180
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 240
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 300
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 360
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 420
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 480
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 540
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 600
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 660
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 720
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 780
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 840
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 900
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 960
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1020
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1080
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1140
nnnagcgcgt tcgccggcaa tgccaacggg ttcaaactgg gaggaaccag cggtcggcaa 1200
tcatcgcatt acccgttgcg tcgcgtttgg caacgtgagc aaaggttttg accagaacaa 1260
taacgcggcg gcgtacggta ataaataaca cgtcatataa aaacggcatt aattatgctt 1320
tcgggagcaa cgtccagtca ggacagaaac attacttccg caacaacgtg tctctgtctg 1380
cttctgcaac ggtcagtaat gcggatgcga gtcaactcat ggatacggac agcggcttcc 1440
gcgtccgatt cgtcagcctg gacacctcgc tggctaccgt atctcgtgat aatgacggca 1500
cgttgccggc aacctcgctg ttccgtctgt cggctaactc aaagctgatt aacgcgggta 1560
cgaaagagaa atcatattag ttacaaggca gcgcaccgga tttgggcgct tttgaacgca 1620
attaagtaat ggaatgttgc gtatcaggct gcaatcgcag ccctgttatt ttgggggtga 1680
aaagattaat caattagtta cagctaccaa cctgtgccca ggaggaatcg ctgcccggaa 1740
cggatgaagt gtaccagttt gcggtataca ggttcccttt gtagacgatc gattggcctg 1800
cttcgttatg agtcgctgcc ggcccgccca gtctttgcta acccagttgg ggtaaacgtt 1860
ggcattgcag caatcagcgg tttgccggtg tgtcggttgt ggtcggttcg tcaacggtgg 1920
tgccatggtg gtatcggttg actgatcggt tgttgtactg gcggcgctgc ccgctttata 1980
tggccagctt tgaatgatcg attttactat tttacccgac tcggtcaggt ttttagagtc 2040
cggataatag gttgatgccc cttcgctttt atcatttaac gcccagtttg cgttgctgat 2100
gttgttgtca cgcatgaacg ttacccaggc gtcggtatct gtctggttca ctccgccatt 2160
gccgtccgcg ttaacggcgc cccactcggt gacgaaaagc gcaataccgt tatttaacgc 2220
ctggcgggct ttagtgcgta atgactcacc atgggttccc gcgtagaaat gcagcgtata 2280
ggcgatattc ttggcgttga ttggatcgcg cgacgcttca tcaacgtttt gcgaccaact 2340
gggcgtaccg acaataatca ggttatccgg gtcaatggcg cgaatggcgg aaatcacggc 2400
ttcggcataa ggtttaatgg tattgctcca tgaaacctga agcggctcgt tgtagatttc 2460
ataaatgaca ttcggcttgt tgccatattt gcgcgccatt tcctggaaga agcgaatggc 2520
ttcactgcga ttgttttctg cagaatgtga gtgccagtca ataatcacat acatatcgtt 2580
ggcgattgcg gcatccacca ctctttcaac tttggccttg ttgccagccg ggtcctgcag 2640
ataaccaccg ctttcctgaa cgcccatagc ggcgcgaaca atgctggatt ccagtctttt 2700
tttcagcgac gcaacggtat cggctgtgta gaatttttcc ccaccccaac cattattact 2760
ccagaataag ctgttgccgg caaaactttt ggctttttca cctgcgtaga ttttattgcc 2820
gctaacggat aacggctcaa cactcgccca ggcattactg gaaaggcagg caaggcttaa 2880
tcctaaacag gcacaagata gaaacagttt tttacgtggg gcgtgttttt ggctatcgat 2940
gactggatgg ttatccgaat aagagagagg cataaatgaa tctccatttc atagactcta 3000
gaggatcacg aaaatcaggc agcgcttaaa tctgccaaaa caccggaggc tgcttcatta 3060
tggcagcgca aggcaatttc agtataaagg gcatcacgat cgaaatcaat gctgccgccc 3120
tgaatggctt ccatcaggtc ttctgcgccg acaacttctg caccagcagc tttagcttct 3180
tcagctttcg ggccacgggc gaaaacaccg acgcgaacat ttttaccggt gcccttcggc 3240
aaggtaacaa caccacgaac catctgatca aggtattaag ccgccggac aggtaaaatg 3300
gttatggcta accgatgccg atattgcgca tcaggatgat accttgaaaa tgttggttgc 3360
tcatgccgaa aaagacggat tgattttca ttccctgatg gcgcgactaa gttgtaacaa 3420
```

```
ttttgcagaa tgggcgctta ttccggcctt cgtctttttc ttccagatgt tgtatccttt 3480
tcggcatatc aataatccca agcataaaac ggccggtgct gccggtggat gcatgttagc 3540
cgatcatcat ctgtgctggt ttggataaac cgttactaat ctggcgcgcg acccgataga 3600
cagcggcttc cccagccccc agtgagctac caatagcaag ggtgccaccc tgctggaaaa 3660
tggagtctag gatttcatta aaactgacgc ttaaagtgaa tttccacatc ccttcaacgg 3720
gcgtttttct ctgaaatatt tttcttatcg gaaatgctcg ctccgtatat tggtgaaaga 3780
gataaatgcc agcgtaacta caggtgacaa aaagcgtgaa ttgcgtcagg caccatataa 3840
acaaaaaata ccccaaaggc atatgaagcc aataaccaat gccacaacct ccggttctga 3900
cgcaggtcgt aataaactca taaatagtga ccagtttaaa gcggttacac agccgcaaca 3960
tcccggttgg accagccgat attcataaaa agtataatca gcatacaaag caaggcatct 4020
ggcttgacct cggcaggcca tcccaatctt gaagtgccta aaatcaagat accgcctaac 4080
cctaccaaca tacctatcgc cgcactaaaa aaatcggctc tgatgcaaaa ggcaaggaca 4140
tcatcaaatt ggttaaaatc tttttcctga aaagcttttg aaccgtaatg cagcaaggtt 4200
tgccatgact gaaaacgggt aatatccgaa atcaggtag caaaagtcgt aatcaaaagc 4260
atgacaccaa acaggttcaa tcccaatgtc tgggatgtcc atgcgacata gacaaaactg 4320
cataccgcat tcaaaacacg gcctgtaatc agaatgccgg tattgccaat gatacgggaa 4380
aagacagatc gaatttgatg ccgtggataa attgccggtt gatcgatccc cgggtaccga 4440
gctcgaattc cgagcttggc gcgcctatgc ggtgtgaaat accgcacaga tgcgtaagga 4500
gaaaataccg catcaggcgc cattcgccat tcaggctgcg caactgttgg gaagggcgat 4560
cggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg gggatgtgc gcaaggcgat 4620
taagttgggt aacgccaggg ttttcccagt cacgacgttg taaaacgacg gccagtgaat 4680
tccatcaacg cttgctgtaa ccaggagcca aagctatgaa tgtacctttt agctactcgt 4740
cacccaccct gagcgttgag gcgttaaagc actctattgc ttataagctg atgtttatca 4800
tcggcaaaga cccggctatc gctaacaagc atgaatggct caacgccacg ctgttcgccg 4860
ttcgcgatcg tatggttgag cgctggctgc gctcaaaccg cgcgcacgtc tctcaggaag 4920
ttcgccaggt ttactacctg tcgatggaat ttttgattgg ccgtacgttg tccaacgcgc 4980
tgttatcgct cggcatttat gaggatgtga acagcgcgct ggaagagatg gggctgaacc 5040
ttgaagaatt aattgatgaa gaaaacgacc cgggcttagg caacggcggt cttggtcgtc 5100
tggcggcctg cttcctcgat tcgcttgcgg cgctggggtt accgggccgc ggctacggta 5160
ttcgctacga ctacgggatg tttaagcaga atatcgtcga tgggcggcag aaagaatccc 5220
cggattactg gctggaatac ggtaacccgt gggagttcga gcgccataat acgcgctaca 5280
aagtgcgctt cggcggacgc attcagcagg aaggtaaata ctcccgctgg gtggagaccg 5340
aagagattat tgccgaagcc tatgaccaga ttatccctgg cttcgacacc gacgccacca 5400
acacgctgcg cctgtggagc gcccaggcca gcagcgagat taacctcggt aaattcaacc 5460
agggcgacta cttcgcggcg gtggaagata aaaaccattc cgagaacgtg tcgcgggtac 5520
tctatccgga tgactcgacc tattcaggac gcgagctgcg cctgcggcag gagtacttcc 5580
tcgtttcggc gacggtgcag gacatcctca gccgccacta ccagctgcat aaaaacctacg 5640
ccaacctggc ggacaaaatc gcgattcatc tcaacgacac gaacccggtg ctgtcgattc 5700
cggagctgat gcgcctgctg attgacgagc ataagatcag ctgggatgag ggnnnnnnnn 5760
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 5820
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 5880
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 5940
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6000
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6060
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6120
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6180
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6240
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6300
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6360
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6420
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6480
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6540
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6600
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6660
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6720
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6780
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6840
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6900
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6960
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7020
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7080
```

```
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7140
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7200
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7260
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7320
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7380
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7440
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7500
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 7560
nnnnnnngaa ttcgagctcg gtacccgggg atcgacataa ccgataggtc ctgcattgat 7620
ggactgaaag gtttcgacgg cggttgtgtg ggttttgctt gcccatctgg cggcatgaat 7680
agtgtcattc atgacgatcc agttcgatat tcaacagacc gtctttgtaa tcggcaccga 7740
caattttgat caataaagcg tttgacctga tgcatgaggg taaatccatt cgttcggttg 7800
ttcttttctg attacctgtc ctgttaacct gtggatatag aaggtcggtt caatgagtag 7860
tattctgacg catctgacaa ttggttccaa tgacctgaag aaggcgcgca tcttttatga 7920
tgctgttttg gaaccgttgg gtatcaaact tattcgcgag gtcgaaggac agcgttttgc 7980
ctatggtaaa gacggcgaag aaggacgcat catcattgta aagcctatta atggtgaagc 8040
cgctaccgct ggaaatggta tcactatcgg tttggcagcg ccttctgatg aagctgtcga 8100
tgcttttat aaagcaggct tggctaatgg cggtaaggat gccggagaac cggggcctcg 8160
tccggctgct aataattctc ggggtgccta tttatatgac cctgaaggca ataaaatctg 8220
cgctttcaat tttaaataag atttctttgg tgcagggtta ttcaaaatag ccctgcattt 8280
tcagtattat agcggccatt atggcttttg ccttgataaa aaatttatca gggctgtttt 8340
tcgtgatgaa tatttttgat ttttcaagaa aagcctgata tcttccaaca tctttctttg 8400
tatataaatg gagcgagcta tggcgcgcgt aactgtcgaa gactgtatcg ataaagttca 8460
taatcgtttc gatttgatcc tctagagtca acctgcttgt tactcgtgat cccattcaca 8520
agggcgaatt aattcgccct tctgttccgt taccaacact tgagccggag gcataatggg 8580
aaagccaatg tggcgttgtt gggcgttgat gctgatggtg tggttcagtg cgtcggctac 8640
ggcggcgaac ggctgggaaa tctataaaag ccgtttcatg accacgacg ggcgcattca 8700
ggataccggc aataagaatg tcagccacac cgaaggtcag ggattcgcca tgctgatggc 8760
ggtgcattac gatgaccgca tcgcgttcga taacctgtgg aactggacgc aaagccacct 8820
gcggaacacg accagcggct tgttctactg gcgttacgat ccgtcggcgg ccaatccggt 8880
ggtggataag aacaacgcct cggatggcga tgtgctgatt gcctgggcgt tgttaaaagc 8940
gggaaataag tggcaggaca accgttacct gcaggcgtcg dacagcatcc agaaagcgat 9000
catcgccagc aatatcattc agtttgcggg ccgcaccgtg atgttgcccg cgcctatgg 9060
tttcaacaag aacagctatg tgatccttaa cccgtcgtat ttcctgttcc cggcctggcg 9120
cgactttgct aaccgcagcc atcttcaggt gtggcggcaa ctgattgacg acagcctgtc 9180
attggtcgga gaaatgcgtt tcggtcaggt cgggctgccg acggactggg cggcgctgaa 9240
cgcggatggc tcgatggcgc cggcgacggc ctgccgtcg cgtttcagtt acgacgccat 9300
tcgtatcccg ctgtatttgt actggtatga cgccaaaacc acggcgctgg tgccgttcca 9360
gctgtactgg cgtaactatc cccgcctgac gacgccggcc tgggttgatg tgctgagcag 9420
taacaccgcg acttacaata tgcagggcgg tttgctggcg gtgcgcgacc tgacgatggg 9480
caacctcgac gggctcagcg atctgccagg cgcatcggaa gattactact cgtcgagcct 9540
gcgcctgctg gtgatgttgg cgcgcggtaa ataaccttat tcttgcggta cacatggcga 9600
ggacgatgtc cttgccattt tccccacttt tatccctctg aatggcgtgt ttttcacgct 9660
ttgttaacct gcttgttact cgtgatccca ttcacaaggg cgaattgacc tgcaggcatg 9720
caagcttggc gtaatcatgg tcatagctgt ttcctgtgtg aaattgttat ccgctcacaa 9780
ttccacacaa catacgagcc ggaagcataa agtgtaaagc ctggggtgcc taatgagtga 9840
gctaactcac attaattgcg ttgcgctcac tgcccgcttt ccagtcggga aacctgtcgt 9900
gccagctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt attgggcgct 9960
cttccgctgg cgcgccaggt cgactctaga ggatccccgg ggaagatctt ccggaagatc 10020
ttcccgagct cgaattaatt ccgcgatgaa ttgatcccgg aagttcctat tctctagaaa 10080
gtataggaac tcgaattggt cgacaagcta gcttgcatgc aagcttgtat tctatagtgt 10140
cacctaaatc gtatgtgtat gatacataag gttatgtatt aattgtagcc gcgttctaac 10200
gacaatatgt acaagcctaa ttgtgtagca tctggcttac tgaagcagac cctatcatct 10260
ctctcgtaaa ctgccgtcag agtcggtttg gttggacgaa ccttctgagt ttctggtaac 10320
gccgttccgc accccggaaa tggtcagcga accaatcagc agggtcatcg ctagcccatg 10380
gctaattctg tcagccgtta agtgttcctg tgtcactgaa aattgctttg agaggctcta 10440
agggcttctc agtgcgttac atccctggct tgttgtccac aaccgttaaa ccttaaaagc 10500
tttaaaagcc ttatatattc tttttttct tataaaactt aaaaccttag aggctattta 10560
agttgctgat ttatattaat tttattgttc aaacatgaga gcttagtacg tgaaacatga 10620
gagcttagta cgttagccat gagagcttag tacgttagcc atgagggttt agttcgttaa 10680
acatgagagc ttagtacgtt aaacatgaga gcttagtacg tgaaacatga gagcttagta 10740
```

```
cgtactatca acaggttgaa ctgcggatct tgcggccgca aaaattaaaa atgaagtttt 10800
gacggtatcg aaccccagag tcccgctcag aagaactcgt caagaaggcg atagaaggcg 10860
atgcgctgcg aatcgggagc ggcgataccg taaagcacga ggaagcggtc agcccattcg 10920
ccgccaagct cttcagcaat atcacgggta gccaacgcta tgtcctgata gcggtccgcc 10980
acacccagcc ggccacagtc gatgaatcca gaaaagcggc cattttccac catgatattc 11040
ggcaagcagg catcgccatg ggtcacgacg agatcctcgc cgtcgggcat ccgcgccttg 11100
agcctggcga acagttcggc tggcgcgagc ccctgatgct cttcgtccag atcatcctga 11160
tcgacaagac cggcttccat ccgagtacgt gctcgctcga tgcgatgttt cgcttggtgg 11220
tcgaatgggc aggtagccgg atcaagcgta tgcagccgcc gcattgcatc agccatgatg 11280
gatactttct cggcaggagc aaggtgagat gacaggagat cctgcccccgg cacttcgccc 11340
aatagcagcc agtcccttcc cgcttcagtg acaacgtcga gcacagctgc gcaaggaacg 11400
cccgtcgtgg ccagccacga tagccgcgct gcctcgtctt ggagttcatt caggcaccg 11460
gacaggtcgg tcttgacaaa aagaaccggg cgcccctgcg ctgacagccg gaacacggcg 11520
gcatcagagc agccgattgt ctgttgtgcc cagtcatagc cgaatagcct ctccacccaa 11580
gcggccggag aacctgcgtg caatccatct tgttcaatca tgcgaaacga tcctcatcct 11640
gtctcttgat ccactagatt attgaagcat ttatcagggt tattgtctca tgagcggata 11700
catatttgaa tgtatttaga aaaataaaca aataggggtt ccgcgcacat ttccccgaaa 11760
agtgccacct gc                                                    11772
```